

Raw Sequence Listing  
Patent Application US/07/599,543E12/17/91  
13:43:49

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Opperman, Hermann  
Ozkaynak, Engin  
Rueger, David C.  
Kuberasampath, Thangavel
- (ii) TITLE OF INVENTION: Osteogenic Proteins
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Testa Hurwitz & Thibeault
  - (B) STREET: 53 State Street
  - (C) CITY: Boston
  - (D) STATE: Massachusetts
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette, 3.50 inch, 720 kb storage
  - (B) COMPUTER: IBM XT
  - (C) OPERATING SYSTEM: DOS 3.30
  - (D) SOFTWARE: ASC II
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/599,543
  - (B) FILING DATE: 18-Oct-90
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 569,920
  - (B) FILING DATE: 20-Aug-90
  - (C) APPLICATION NUMBER: US 315,342
  - (D) FILING DATE: 23-Feb-89
  - (E) APPLICATION NUMBER: US 422,699
  - (F) FILING DATE: 17-Oct-89

Patent Application US/07/599,543E

54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69 (2) INFORMATION FOR SEQ ID NO:1:  
70 (i) SEQUENCE CHARACTERISTICS:  
71 (A) LENGTH: 139 amino acids  
72 (B) TYPE: amino acid  
73 (D) TOPOLOGY: linear  
74 (ii) MOLECULE TYPE: protein  
75 (ix) FEATURE:  
76 (A) NAME: mOP2 (mature)  
77 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
78  
79 Ala Ala Arg Pro Leu Lys Arg Arg Gln  
80 1 5  
81 Pro Lys Lys Thr Asn Glu Leu Pro His  
82 10 15  
83 Pro Asn Lys Leu Pro Gly Ile Phe Asp  
84 20 25  
85 Asp Gly His Gly Ser Arg Gly Arg Glu  
86 30 35  
87 Val Cys Arg Arg His Glu Leu Tyr Val  
88 40 45  
89 Arg Phe Arg Asp Leu Gly Trp Leu Asp  
90 50  
91 Trp Val Ile Ala Pro Gln Gly Tyr Ser  
92 55 60  
93 Ala Tyr Tyr Cys Glu Gly Glu Cys Ala  
94 65 70  
95 Phe Pro Leu Asp Ser Cys Met Asn Ala  
96 75 80  
97 Thr Asn His Ala Ile Leu Gln Ser Leu  
98 85 90  
99 Val His Leu Met Lys Pro Asp Val Val  
100 95  
101 Pro Lys Ala Cys Cys Ala Pro Thr Lys  
102 100 105  
103 Leu Ser Ala Thr Ser Val Leu Tyr Tyr  
104 110 115  
105  
106

107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135 Asp Ser Ser Asn Asn Val Ile Leu Arg  
136 120 125  
137 Lys His Arg Asn Met Val Val Lys Ala  
138 130 135  
139 Cys Gly Cys His  
140  
141

142 (2) INFORMATION FOR SEQ ID NO:2:  
143 (i) SEQUENCE CHARACTERISTICS:  
144 (A) LENGTH: 1930 base pairs  
145 (B) TYPE: nucleic acid  
146 (C) STRANDEDNESS: single  
147 (D) TOPOLOGY: linear  
148 (ii) MOLECULE TYPE: cDNA  
149 (iii) HYPOTHETICAL: no  
150 (iv) ANTI-SENSE: no  
151 (vi) ORIGINAL SOURCE:  
152 (A) ORGANISM: mouse  
153 (F) TISSUE TYPE: embryo  
154 (ix) FEATURE:  
155 (A) NAME: mOP2  
156 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

157  
158 GGAATTCCGC TGCCAGGCAC AGGTGCGCCG TCTGGTCCTC 40  
159 CCCGTCTGGC GTCAGCCGAG CCCGACCAGC TACCAGTGGA 80

## Patent Application US/07/599,543E

```
160  TCGCGCCCGG CTGAAAGTCC GAG ATG GCT ATG CGT      115
161                      Met Ala Met Arg
162                      1
163  CCC GGG CCA CTC TGG CTA TTG GGC CTT GCT CTG      148
164  Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu
165    5          10          15
166  TGC GCG CTG GGA GGC GGC CAC GGT CCC GGT CCC      181
167  Cys Ala Leu Gly Gly Gly His Gly Pro Gly Pro
168          20          25
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200  CCG CAC ACC TGT CCC CAG CGT CGC CTG GGA GCG      214
201  Pro His Thr Cys Pro Gln Arg Arg Leu Gly Ala
202          30          35
203  CGC GAC CGG GAC ATG CAG CGT GAA ATC CTG CCG      247
204  Arg Asp Arg Asp Met Gln Arg Glu Ile Leu Pro
205    40          45
206  GTG CTC GGG CTA CCG GGA CGC CCC GAC CCC GTG      280
207  Val Leu Gly Leu Pro Gly Arg Pro Asp Pro Val
208    50          55
209  CAC AAC CCG CCG CTG CCC GGC ACG CAG CGT GCG      313
210  His Asn Pro Pro Leu Pro Gly Thr Gln Arg Ala
211    60          65          70
212  CCC CTC TTC ATG TTG GAC CTA TAC CAC GCC ATG      346
```

213 Pro Leu Phe Met Leu Asp Leu Tyr His Ala Met  
214 75 80  
215 ACC GAT GAC GAC GAC GGC GGG CCA CCA CAG GCT 379  
216 Thr Asp Asp Asp Asp Gly Gly Pro Pro Gln Ala  
217 85 90  
218 CAC TTA GGC CGT GCC GAC CTG GTC ATG AGC TTC 412  
219 His Leu Gly Arg Ala Asp Leu Val Met Ser Phe  
220 95 100  
221 GTC AAC ATG GTG GAA CGC GAC CGT ACC CTG GGC 445  
222 Val Asn Met Val Glu Arg Asp Arg Thr Leu Gly  
223 105 110  
224 TAC CAG GAG CCA CAC TGG AAG GAA TTC CAC TTT 478  
225 Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe  
226 115 120 125  
227 GAC CTA ACC CAG ATC CCT GCT GGG GAG GCT GTC 511  
228 Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val  
229 130 135  
230 ACA GCT GCT GAG TTC CGG ATC TAC AAA GAA CCC 544  
231 Thr Ala Ala Glu Phe Arg Ile Tyr Lys Glu Pro  
232 140 145  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265 AGC ACC CAC CCG CTC AAC ACA ACC CTC CAC ATC 577

## Patent Application US/07/599,543E

266 Ser Thr His Pro Leu Asn Thr Thr Leu His Ile  
267 150 155  
268 AGC ATG TTC GAA GTG GTC CAA GAG CAC TCC AAC 610  
269 Ser Met Phe Glu Val Val Gln Glu His Ser Asn  
270 160 165  
271 AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG 643  
272 Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln  
273 170 175 180  
274 ACG CTC CGA TCT GGG GAC GAG GGC TGG CTG GTG 676  
275 Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu Val  
276 185 190  
277 CTG GAC ATC ACA GCA GCC AGT GAC CGA TGG CTG 709  
278 Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu  
279 195 200  
280 CTG AAC CAT CAC AAG GAC CTG GGA CTC CGC CTC 742  
281 Leu Asn His His Lys Asp Leu Gly Leu Arg Leu  
282 205 210  
283 TAT GTG GAA ACC GCG GAT GGG CAC AGC ATG GAT 775  
284 Tyr Val Glu Thr Ala Asp Gly His Ser Met Asp  
285 215 220  
286 CCT GGC CTG GCT GGT CTG CTT GGA CGA CAA GCA 808  
287 Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala  
288 225 230 235  
289 CCA CGC TCC AGA CAG CCT TTC ATG GTA ACC TTC 841  
290 Pro Arg Ser Arg Gln Pro Phe Met Val Thr Phe  
291 240 245  
292 TTC AGG GCC AGC CAG AGT CCT GTG CGG GCC CCT 874  
293 Phe Arg Ala Ser Gln Ser Pro Val Arg Ala Pro  
294 250 255  
295 CGG GCA GCG AGA CCA CTG AAG AGG AGG CAG CCA 907  
296 Arg Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro  
297 260 265  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318

319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330 AAG AAA ACG AAC GAG CTT CCG CAC CCC AAC AAA 940  
331 Lys Lys Thr Asn Glu Leu Pro His Pro Asn Lys  
332 270 275  
333 CTC CCA GGG ATC TTT GAT GAT GGC CAC GGT TCC 973  
334 Leu Pro Gly Ile Phe Asp Asp Gly His Gly Ser  
335 280 285 290  
336 CGC GGC AGA GAG GTT TGC CGC AGG CAT GAG CTC 1006  
337 Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu  
338 295 300  
339 TAC GTC AGA TTC CGT GAC CTT GGC TGG CTG GAC 1039  
340 Tyr Val Arg Phe Arg Asp Leu Gly Trp Leu Asp  
341 305 310  
342 TGG GTC ATC GCC CCC CAG GGC TAC TCT GCC TAT 1072  
343 Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr  
344 315 320  
345 TAC TGT GAG GGG GAG TGT GCT TTC CCA CTG GAC 1105  
346 Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp  
347 325 330  
348 TCC TGT ATG AAC GCC ACC AAC CAT GCC ATC TTG 1138  
349 Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu  
350 335 340 345  
351 CAG TCT CTG GTG CAC CTG ATG AAG CCA GAT GTT 1171  
352 Gln Ser Leu Val His Leu Met Lys Pro Asp Val  
353 350 355  
354 GTC CCC AAG GCA TGC TGT GCA CCC ACC AAA CTG 1204  
355 Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu  
356 360 365  
357 AGT GCC ACC TCT GTG CTG TAC TAT GAC AGC AGC 1237  
358 Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser  
359 370 375  
360 AAC AAT GTC ATC CTG CGT AAA CAC CGT AAC ATG 1270  
361 Asn Asn Val Ile Leu Arg Lys His Arg Asn Met  
362 380 385  
363  
364  
365  
366  
367  
368  
369  
370  
371

372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395 GTG GTC AAG GCC TGT GGC TGC CAC 1294  
396 Val Val Lys Ala Cys Gly Cys His  
397 390 400  
398 TGAGGCCCGG CCCAGCATCC TGCTTCTACT ACCTTACCAT 1334  
399 CTGGCCGGGG CCCTCTCCAG AGGCAGAAAC CCTTCTATGT 1374  
400 TATCATAGCT CAGACAGGGG CAATGGGAGG CCCTTCACTT 1414  
401 CCCCTGGCCA CTTCTGCTA AAATTCTGGT CTTTCCCAGT 1454  
402 TCCTCTGTCC TTCATGGGGT TTCGGGGCTA TCACCCCGCC 1494  
403 CTCTCCATCC TCCTACCCCA AGCATAGACT GAATGCACAC 1534  
404 AGCATCCCAG AGCTATGCTA ACTGAGAGGT CTGGGGTCAG 1574  
405 CACTGAAGGC CCACATGAGG AAGACTGATC CTTGGCCATC 1614  
406 CTCAGCCAC AATGGCAAAT TCTGGATGGT CTAAGAAGCC 1654  
407 CTGGAATTCT AAACATAGATG ATCTGGGCTC TCTGCACCAT 1694  
408 TCATTGTGGC AGTTGGGACA TTTTATAGGTA TAACAGACAC 1734  
409 ATACACTTAG ATCAATGCAT CGCTGTACTC CTTGAAATCA 1774  
410 GAGCTAGCTT GTTAGAAAA GAATCAGAGC CAGGTATAGC 1814  
411 GGTGCATGTC ATTAATCCCA GCGCTAAAGA GACAGAGACA 1854  
412 GGAGAATCTC TGTGAGTTCA AGGCCACATA GAAAGAGCCT 1894  
413 GTCTCGGGAG CAGGAAAAA AAAAAAACG GAATTC 1930  
414  
415

416 (2) INFORMATION FOR SEQ ID NO:3:  
417 (i) SEQUENCE CHARACTERISTICS:  
418 (A) LENGTH: 139 amino acids  
419 (B) TYPE: amino acid  
420 (D) TOPOLOGY: linear  
421 (ii) MOLECULE TYPE: protein  
422 (ix) FEATURE:  
423 (A) NAME: hOP2 (mature)  
424 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:



Raw Sequence Listing  
Patent Application US/07/599,543E12/17/91  
13:44:21

```
425
426 Ala Val Arg Pro Leu Arg Arg Arg Gln
427 1 5
428 Pro Lys Lys Ser Asn Glu Leu Pro Gln
429 10 15
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462 Ala Asn Arg Leu Pro Gly Ile Phe Asp
463 20 25
464 Asp Val His Gly Ser His Gly Arg Gln
465 30 35
466 Val Cys Arg Arg His Glu Leu Tyr Val
467 40 45
468 Ser Phe Gln Asp Leu Gly Trp Leu Asp
469 50
470 Trp Val Ile Ala Pro Gln Gly Tyr Ser
471 55 60
472 Ala Tyr Tyr Cys Glu Gly Glu Cys Ser
473 65 70
474 Phe Pro Leu Asp Ser Cys Met Asn Ala
475 75 80
476 Thr Asn His Ala Ile Leu Gln Ser Leu
477 85 90
```

478 Val His Leu Met Lys Pro Asn Ala Val  
479 95  
480 Pro Lys Ala Cys Cys Ala Pro Thr Lys  
481 100 105  
482 Leu Ser Ala Thr Ser Val Leu Tyr Tyr  
483 110 115  
484 Asp Ser Ser Asn Asn Val Ile Leu Arg  
485 120 125  
486 Lys Ala Arg Asn Met Val Val Lys Ala  
487 130 135  
488 Cys Gly Cys His  
489  
490

491 (2) INFORMATION FOR SEQ ID NO:4:  
492 (i) SEQUENCE CHARACTERISTICS:  
493 (A) LENGTH: 1941 base pairs  
494 (B) TYPE: nucleic acid  
495 (C) STRANDEDNESS: single  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526

527 (D) TOPOLOGY: linear  
528 (ii) MOLECULE TYPE: cDNA  
529 (iii) HYPOTHETICAL: no  
530 (iv) ANTI-SENSE: no

## Patent Application US/07/599,543E

531 (vi) ORIGINAL SOURCE:  
532 (A) ORGANISM: homo sapiens  
533 (F) TISSUE TYPE: hippocampus  
534 (ix) FEATURE:  
535 (A) NAME: hOP2  
536 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

537  
538 GGAATTCCGG CCACAGTGGC GCCGGCAGAG CAGGAGTGGC 40  
539 TGGAGGAGCT GTGGTTGGAG CAGGAGGTGG CACGGCAGGG 80  
540 CTGGAGGGCT CCCTATGAGT GGCGGAGACG GCCCAGGAGG 120  
541 CGCTGGAGCA ACAGCTCCCA CACCGCACCA AGCGGTGGCT 160  
542 GCAGGAGCTC GCCCATCGCC CCTGCGCTGC TCGGACCGCG 200  
543 GCCACAGCCG GACTGGCGGG TACGGCGGCG ACAGAGGCAT 240  
544 TGGCCGAGAG TCCCAGTCCG CAGAGTAGCC CCGGCCTCGA 280  
545 GGCGGTGGCG TCCCAGTCCCT CTCCGTCCAG GAGCCAGGAC 320  
546 AGGTGTCGCG CGGCGGGGCT CCAGGGACCG CGCCTGAGGC 360  
547 CGGCTGCCCC CGCGTCCCGC CCCGCCCCGC CGCCCGCCGC 400  
548 CCGCCGAGCC CAGCCTCCTT GCCGTCGGGG CGTCCCCAGG 440  
549 CCCTGGGTCG GCCGCGGAGC CGATGCGCGC CCGCTGAGCG 480  
550 CCCCAGCTGA GCGCCCCCGG CCTGCC ATG ACC GCG CTC 518  
551 Met Thr Ala Leu  
552 1  
553 CCC GGC CCG CTC TGG CTC CTG GGC CTG GCG CTA 551  
554 Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu  
555 5 10 15  
556 TGC GCG CTG GGC GGG GGC GGC CCC GGC CTG CGA 584  
557 Cys Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg  
558 20 25  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583

## Patent Application US/07/599,543E

584  
585  
586  
587  
588  
589  
590  
591  
592 CCC CCG CCC GGC TGT CCC CAG CGA CGT CTG GGC 617  
593 Pro Pro Pro Gly Cys Pro Gly arg Arg Leu Gly  
594 30 35  
595 GCG CGC GAC CGG GAC GTG CAG CGC GAG ATC CTG 650  
596 Ala Arg Asp Arg Asp Val Gln Arg Glu Ile Leu  
597 40 45  
598 GCG GTG CTC GGG CTG CCT GGG CGG CCC CGG CCC 683  
599 Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro  
600 50 55  
601 CGC GCG CCA CCC GCC GCC TCC CGG CTG CCC GCG 716  
602 Arg Ala Pro Pro Ala Ala Ser Arg Leu Pro Ala  
603 60 65 70  
604 TCC GCG CCG CTC TTC ATG CTG GAC CTG TAC CAC 749  
605 Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr His  
606 75 80  
607 CGC ATG GCC GGC GAC GAC GAC GAG GAC GGC GCC 782  
608 Arg Met Ala Gly Asp Asp Asp Glu Asp Gly Ala  
609 85 90  
610 GCG GAG GCC CTG GGC CGC GCC GAC CTG GTC ATG 815  
611 Ala Glu Ala Leu Gly Arg Ala Asp Leu Val Met  
612 95 100  
613 AGC TTC GTT AAC ATG GTG GAG CGA GAC CGT GCC 848  
614 Ser Phe Val Asn Met Val Glu Arg Asp Arg Ala  
615 105 110  
616 CTG GGC CAC CAG GAG CCC CAT TGG AAG GAG TTC 881  
617 Leu Gly His Gln Glu Pro His Trp Lys Glu Phe  
618 115 120 125  
619 CGC TTT GAC CTG ACC CAG ATC CCG GCT GGG GAG 914  
620 Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu  
621 130 135  
622 GCG GTC ACA GCT GCG GAG TTC CGG ATT TAC AAG 947  
623 Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys  
624 140 145  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636

637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657 GTG CCC AGC ATC CAC CTG CTC AAC AGG ACC CTC 980  
658 Val Pro Ser Ile His Leu Leu Asn Arg Thr Leu  
659 150 155  
660 CAC GTC AGC ATG TTC CAG GTG GTC CAG GAG CAG 1013  
661 His Val Ser Met Phe Gln Val Val Gln Glu Gln  
662 160 165  
663 TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT 1046  
664 Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp  
665 170 175 180  
666 CTT CAG ACG CTC CGA GCT GGA GAC GAG GGC TGG 1079  
667 Leu Gln Thr Leu Arg Ala Gly Asp Glu Gly Trp  
668 185 190  
669 CTG GTG CTG GAT GTC ACA GCA GCC AGT GAC TGC 1112  
670 Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys  
671 195 200  
672 TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA CTC 1145  
673 Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu  
674 205 210  
675 CGC CTC TAT GTG GAG ACT GAG GAC GGG CAC AGC 1178  
676 Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser  
677 215 220  
678 GTG GAT CCT GGC CTG GCC GGC CTG CTG GGT CAA 1211  
679 Val Asp Pro Gly Leu Ala Gly Leu Leu Gly Gln  
680 225 230 235  
681 CGG GCC CCA CGC TCC CAA CAG CCT TTC GTG GTC 1244  
682 Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val  
683 240 245  
684 ACT TTC TTC AGG GCC AGT CCG AGT CCC ATC CGC 1277  
685 Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg  
686 250 255  
687 ACC CCT CGG GCA GTG AGG CCA CTG AGG AGG AGG 1310  
688 Thr Pro Arg Ala Val Arg Pro Leu Arg Arg Arg  
689 260 265

690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722 CAG CCG AAG AAA AGC AAC GAG CTG CCG CAG GCC 1343  
723 Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala  
724 270 275  
725 AAC CGA CTC CCA GGG ATC TTT GAT GAC GTC CAC 1376  
726 Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His  
727 280 285 290  
728 GGC TCC CAC GGC CGG CAG GTC TGC CGT CGG CAC 1409  
729 Gly Ser His Gly Arg Gln Val Cys Arg Arg His  
730 295 300  
731 GAG CTC TAC GTC AGC TTC CAG GAC CTC GGC TGG 1442  
732 Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp  
733 305 310  
734 CTG GAC TGG GTC ATC GCT CCC CAA GGC TAC TCG 1475  
735 Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser  
736 315 320  
737 GCC TAT TAC TGT GAG GGG GAG TGC TCC TTC CCA 1508  
738 Ala Tyr Tyr Cys Glu Gly Glu Cys Ser Phe Pro  
739 325 330  
740 CTG GAC TCC TGC ATG AAT GCC ACC AAC CAC GCC 1541  
741 Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala  
742 335 340 345

## Patent Application US/07/599,543E

743 ATC CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA 1574  
744 Ile Leu Gln Ser Leu Val His Leu Met Lys Pro  
745 350 355  
746 AAC GCA GTC CCC AAG GCG TGC TGT GCA CCC ACC 1607  
747 Asn Ala Val Pro Lys Ala Cys Cys Ala Pro Thr  
748 360 365  
749 AAG CTG AGC GCC ACC TCT GTG CTC TAC TAT GAC 1640  
750 Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp  
751 370 375  
752 AGC AGC AAC AAC GTC ATC CTG CGC AAA GCC CGC 1673  
753 Ser Ser Asn Asn Val Ile Leu Arg Lys Ala Arg  
754 380 385  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787 AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC 1703  
788 Asn Met Val Val Lys Ala Cys Gly Cys His  
789 390 395  
790 TGAGTCAGCC CGCCCAGCCC TACTGCAGCA ATTCACTGGC 1743  
791 CGTCGTTTTA CAACGTGTGA CTGGGAAAAC CCTGGCGTTA 1783  
792 CCCAACTTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG 1823  
793 CTGGCTAATA GCGAAGAGGC CCCGCACCGA TCGCCCTTCC 1863  
794 CAACAGTTGC GCCCCAGTGA ATGGCGAATG GCAAATTGTA 1903  
795 AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTT 1941

796  
797  
798 (2) INFORMATION FOR SEQ ID NO:5:  
799 (i) SEQUENCE CHARACTERISTICS:  
800 (A) LENGTH: 98 amino acids  
801 (B) TYPE: amino acid  
802 (D) TOPOLOGY: linear  
803 (ii) MOLECULE TYPE: protein  
804 (ix) FEATURE:  
805 (D) OTHER INFORMATION: wherein "res."  
806 means "residue" and Xaa at res. 2 = (Lys or Arg); Xaa at  
807 res.3 = (Lys or Arg); Xaa res.9 = (Ser or Arg); Xaa at  
808 res.11 = (Arg or Gln); Xaa at res.16 = (Gln or Leu); Xaa  
809 at res. 19 = (Ile or Val); Xaa at res.23 = (Glu or Gln);  
810 Xaa at res.26 = (Ala or Ser); Xaa at res. 34 = (Ala or  
811 or Ser); Xaa at res.38= (Asn or Asp); Xaa at res. 40 =  
812 (Tyr or Cys); Xaa at res.49 = (Val or Leu); Xaa at  
813 res.52= (His or Asn); Xaa at res. 53 = (Phe or  
814 Leu); Xaa at res. 54 = (Ile or Met); Xaa at res. 55 = (Asn  
815 or Lys); Xaa at res. 56 = (Glu, Asp or Asn); Xaa at res.  
816 57=(Thr, Ala or Val); Xaa at res. 61 = (Pro or Ala);  
817 Xaa at res. 67=(gln or Lys); Xaa at res. 69 =  
818 (Asn or Ser); Xaa at 71=(Ile or Thr); Xaa at res.  
819 76= (Phe or Tyr); Xaa at res. 78 = (Asp, Glu or Ser);  
820 Xaa at res. 80= (Ser or Asn); Xaa at res. 84 = (Ile or  
821 Asp); Xaa at res. 85 Arg); Xaa at res. 87 = (Tyr, Ala  
822 or His); and Xaa at res. 93=(Arg or Lys)

823  
824 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:  
825  
826 Cys Xaa Xaa His Glu Leu Tyr Val Xaa Phe  
827 1 5 10  
828 Xaa Asp Leu Gly Trp Xaa Asp Trp Xaa Ile  
829 15 20  
830 Ala Pro Xaa Gly Tyr Xaa Ala Tyr Tyr Cys  
831 25 30  
832 Glu Gly Cys Xaa Phe Pro Leu Xaa Ser Xaa  
833 35 40  
834 Met Asn Ala Thr Asn His Ala Ile Xaa Thr  
835 45 50  
836 Leu Xaa Xaa Xaa Xaa Xaa Xaa Val  
837 55  
838 Pro Lys Xaa Cys Cys Ala Pro Thr Xaa Leu  
839 60 65  
840  
841  
842  
843  
844  
845  
846  
847  
848



849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901

Xaa Ala Xaa Ser Val Leu Tyr Xaa Asp  
70 75  
Xaa Ser Xaa Asn Val Xaa Leu Xaa Lys  
80 85  
Xaa Pro Asn Met Val Val Xaa Ala Cys Gly  
90 95  
Cys His

- (2) INFORMATION FOR SEQ ID NO:6:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 437 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: no  
(iv) ANTI-SENSE: no  
(vi) ORIGINAL SOURCE:

902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919 (A) ORGANISM: Human  
920 (F) TISSUE TYPE: placenta  
921 (ix) FEATURE:  
922 (A) NAME: OP1  
923 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:  
924  
925 TCC ACG GGG 9  
926 Ser Thr Gly  
927 1  
928 AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG 42  
929 Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr  
930 5 10  
931 CCC AAG AAC CAG GAA GCC CTG CGG ATG GCC AAC 75  
932 Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn  
933 15 20 25  
934 GTG GCA GAG AAC AGC AGC AGC GAC CAG AGG CAG 108  
935 Val Ala Glu Asn Ser Ser Ser Asp Gln Arg Gln  
936 30 35  
937 GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC 141  
938 Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe  
939 40 45  
940 CGA GAC CTG GGC TGG CAG GAC TGG ATC ATC GCG 174  
941 Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala  
942 50 55  
943 CCT GAA GGC TAC GCC GCC TAC TAC TGT GAG GGG 207  
944 Pro Glu Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly  
945 60 65  
946 GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG AAC 240  
947 Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn  
948 70 75 80  
949 GCC ACC AAC CAC GCC ATC GTG CAG ACG CTG GTC 273  
950 Ala Thr Asn His Ala Ile Val Gln Thr Leu Val  
951 85 90  
952  
953  
954

955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984 CAC TTC ATC AAC CCG GAA ACG GTG CCC AAG CCC 306  
985 His Phe Ile Asn Pro Glu Thr Val Pro Lys Pro  
986 95 100  
987 TGC TGT GCG CCC ACG CAG CTC AAT GCC ATC TCC 339  
988 Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser  
989 105 110  
990 GTC CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC 372  
991 Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile  
992 115 120  
993 CTG AAG AAA TAC AGA AAC ATG GTG GTC CGG GCC 405  
994 Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala  
995 125 130 135  
996 TGT GGC TGC CAC TAGCTCCTCC GAGAATTCAG 437  
997 Cys Gly Cys His  
998  
999

1000 (2) INFORMATION FOR SEQ ID NO:7:  
1001 (i) SEQUENCE CHARACTERISTICS:  
1002 (A) LENGTH: 102 amino acids  
1003 (B) TYPE: amino acid  
1004 (D) TOPOLOGY: linear  
1005 (ii) MOLECULE TYPE: protein  
1006 (ix) FEATURE:  
1007 (D) OTHER INFORMATION:

1008 wherein each Xaa independently represents one of  
1009 the 20 naturally occurring L-isomer, a-amino acids.

1010  
1011 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

1012  
1013 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1014 1 5 10  
1015 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1016 15 20  
1017 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa  
1018 25 30  
1019 Cys Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa  
1020 35 40  
1021 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1022 45 50 55

1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053

1054 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys  
1055 60 65  
1056 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1057 70 75  
1058 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1059 80 85  
1060 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys

1061           90           95  
1062       Xaa Cys Xaa  
1063       100  
1064  
1065 (2) INFORMATION FOR SEQ ID NO:8:  
1066       (i) SEQUENCE CHARACTERISTICS:  
1067           (A) LENGTH: 97 amino acids  
1068           (B) TYPE: amino acid  
1069           (D) TOPOLOGY: linear  
1070       (ii) MOLECULE TYPE: protein  
1071       (ix) FEATURE:  
1072       (D) OTHER INFORMATION:  
1073 wherein each Xaa independently represents one of  
1074 the 20 naturally occurring L-isomer, α-amino acids  
1075  
1076       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:  
1077  
1078       Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1079           1           5           10  
1080       Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1081           15           20  
1082       Xaa Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa  
1083           25           30  
1084       Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1085           35           40  
1086       Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1087           45           50           55  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113

1114  
1115  
1116  
1117  
1118  
1119  
1120 Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Xaa Xaa Xaa  
1121 60 65  
1122 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1123 70 75  
1124 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
1125 80 85  
1126 Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa  
1127 90 95  
1128  
1129

1130 (2) INFORMATION FOR SEQ ID NO:9:  
1131 (i) SEQUENCE CHARACTERISTICS:  
1132 (A) LENGTH: 136 amino acids  
1133 (B) TYPE: amino acid  
1134 (D) TOPOLOGY: linear  
1135 (ii) MOLECULE TYPE: protein  
1136 (ix) FEATURE:  
1137 (A) NAME: hOP-2P  
1138 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:  
1139

1140 Pro Leu Arg Arg Arg Gln  
1141 1 5  
1142 Pro Lys Lys Ser Asn Glu Leu Pro Gln  
1143 10 15  
1144 Ala Asn Arg Leu Pro Gly Ile Phe Asp  
1145 20  
1146 Asp Val Asn Gly Ser His Gly Arg Gln  
1147 25 30  
1148 Val Cys Arg Arg His Glu Leu Tyr Val  
1149 35 40  
1150 Ser Phe Gln Asp Leu Gly Trp Leu Asp  
1151 45 50  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166

1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184 Tyr Val Ile Ala Pro Gln Gly Tyr Ser  
1185 55 60  
1186 Ala Tyr Tyr Cys Glu Gly Glu Cys Ser  
1187 65  
1188 Phe Pro Leu Asp Ser Cys Met Asn Ala  
1189 70 75  
1190 Thr Asn His Ala Ile Leu Gln Ser Leu  
1191 80 85  
1192 Val His Leu Met Lys Pro Asn Ala Val  
1193 90 95  
1194 Pro Lys Ala Cys Cys Ala Pro Thr Lys  
1195 100 105  
1196 Leu Ser Ala Thr Ser Val Leu Tyr Tyr  
1197 110  
1198 Asp Glu Ser Asn Asn Val Ile Leu Arg  
1199 115 120  
1200 Lys Ala Arg Asn Met Val Val Lys Ala  
1201 125 130  
1202 Cys Gly Cys His  
1203 135  
1204  
1205  
1206 (2) INFORMATION FOR SEQ ID NO:10:  
1207 (i) SEQUENCE CHARACTERISTICS:  
1208 (A) LENGTH: 133 amino acids  
1209 (B) TYPE: amino acid  
1210 (D) TOPOLOGY: linear  
1211 (ii) MOLECULE TYPE: protein  
1212 (ix) FEATURE:  
1213 (A) NAME: hOP-2R  
1214 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:  
1215  
1216 Arg Arg Gln  
1217 1  
1218 Pro Lys Lys Ser Asn Glu Leu Pro Gln  
1219 5 10

1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251 Ala Asn Arg Leu Pro Gly Ile Phe Asp  
1252 15 20  
1253 Asp Val Asn Gly Ser His Gly Arg Gln  
1254 25 30  
1255 Val Cys Arg Arg His Glu Leu Tyr Val  
1256 35  
1257 Ser Phe Gln Asp Leu Gly Trp Leu Asp  
1258 40 45  
1259 Tyr Val Ile Ala Pro Gln Gly Tyr Ser  
1260 50 55  
1261 Ala Tyr Tyr Cys Glu Gly Glu Cys Ser  
1262 60 65  
1263 Phe Pro Leu Asp Ser Cys Met Asn Ala  
1264 70 75  
1265 Thr Asn His Ala Ile Leu Gln Ser Leu  
1266 80  
1267 Val His Leu Met Lys Pro Asn Ala Val  
1268 85 90  
1269 Pro Lys Ala Cys Cys Ala Pro Thr Lys  
1270 95 100  
1271 Leu Ser Ala Thr Ser Val Leu Tyr Tyr  
1272 105 110



1273 Asp Glu Ser Asn Asn Val Ile Leu Arg  
1274 115 120  
1275 Lys Ala Arg Asn Met Val Val Lys Ala  
1276 125  
1277 Cys Gly Cys His  
1278 130  
1279  
1280

1281 (2) INFORMATION FOR SEQ ID NO:11:

1282 (i) SEQUENCE CHARACTERISTICS:

1283 (A) LENGTH: 160 amino acids

1284 (B) TYPE: amino acid  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314

1315 (D) TOPOLOGY: linear

1316 (ii) MOLECULE TYPE: protein

1317 (ix) FEATURE:

1318 (A) NAME: hOP-2S

1319 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  
1320

1321 Ser Gln Gln

1322 1

1323 Pro Phe Val Val Thr Phe Phe Arg Ala

1324 5 10

1325 Ser Pro Ser Pro Ile Arg Thr Pro Arg



Patent Application US/07/099,543E

1379  
1380  
1381  
1382 Leu Ser Ala Thr Ser Val Leu Tyr Tyr  
1383 130 135  
1384 Asp Glu Ser Asn Asn Val Ile Leu Arg  
1385 140 145  
1386 Lys Ala Arg Asn Met Val Val Lys Ala  
1387 150 155  
1388 Cys Gly Cys His  
1389 160

## LINE ERROR

## ORIGINAL TEXT

24 Wrong application Serial Number  
809 Response Exceeds Line Limitations  
810 Response Exceeds Line Limitations  
811 Response Exceeds Line Limitations  
812 Response Exceeds Line Limitations  
813 Response Exceeds Line Limitations  
814 Response Exceeds Line Limitations  
815 Response Exceeds Line Limitations  
816 Response Exceeds Line Limitations  
817 Response Exceeds Line Limitations  
818 Response Exceeds Line Limitations  
819 Response Exceeds Line Limitations  
820 Response Exceeds Line Limitations  
821 Response Exceeds Line Limitations  
822 Response Exceeds Line Limitations

(A) APPLICATION NUMBER: US 07/599,543  
at res. 19 = (Ile or Val); Xaa at res.23  
Xaa at res.26 = (Ala or Ser); Xaa at res  
or Ser); Xaa at res.38= (Asn or Asp); Xa  
(Tyr or Cys); Xaa at res.49 = (Val or Le  
res.52= (His or Asn); Xaa at res. 53 = (  
Leu); Xaa at res. 54 = (Ile or Met); Xaa  
or Lys); Xaa at res. 56 = (Glu, Asp or A  
57=(Thr, Ala or Val); Xaa at res. 61 = (  
Xaa at res. 67=(gln or Lys); Xaa at res.  
(Asn or Ser); Xaa at 71=(Ile or Thr); Xa  
76= (Phe or Tyr); Xaa at res. 78 = (Asp,  
Xaa at res. 80= (Ser or Asn); Xaa at res  
Asp); Xaa at res. 85 Arg); Xaa at res. 8  
or His); and Xaa at res. 93=(Arg or Lys)

PAGE: 1

SEQUENCE MISSING ITEM REPORT  
PATENT APPLICATION US/07/5 543E

DATE: 12/17/91  
TIME: 13:46:34

MANDATORY IDENTIFIER THAT WAS NOT FOUND

## LINE ORIGINAL TEXT

## CORRECTED TEXT

76 (A) NAME: mOP2 (mature)  
155 (A) NAME: mOP2  
423 (A) NAME: hOP2 (mature)  
535 (A) NAME: hOP2  
922 (A) NAME: OP1  
1137 (A) NAME: hOP-2P  
1213 (A) NAME: hOP-2R  
1318 (A) NAME: hOP-2S

(A) NAME/KEY: mOP2 (mature)  
(A) NAME/KEY: mOP2  
(A) NAME/KEY: hOP2 (mature)  
(A) NAME/KEY: hOP2  
(A) NAME/KEY: OP1  
(A) NAME/KEY: hOP-2P  
(A) NAME/KEY: hOP-2R  
(A) NAME/KEY: hOP-2S